

## Major

1. **Unclear attribution of model biases to microphysics vs. dynamics — needs stronger evidence and clearer separation.**

Explanation: the manuscript frequently attributes the LWC and  $N_c$  biases both to thermodynamic/dynamical biases (e.g., wet/cold/TKE biases) and to structural microphysics limits (single-mode gamma DSD). These arguments appear repeatedly (e.g. discussion of LWP/LWC overestimate and TKE link). The causal chain is plausible but not demonstrated strongly enough: is excess condensation producing high  $N_c$ , or is excessive activation due to turbulence producing excess LWC? See discussion and budget summaries.

add (or more clearly present existing) quantitative diagnostics that separate dynamical from microphysical causes. Show time/height cross-correlations (or regression) between modeled TKE perturbations and activation rate /  $N_c$  production to support the "TKE  $\rightarrow$  activation  $\rightarrow N_c \rightarrow$  LWC" chain. (Add a short figure or table with correlation coefficients or a simple sensitivity test.)

If available, present a sensitivity experiment (or at least a quantitative estimate) where only microphysics is changed (e.g., change activation parameter) and separately where only dynamics (e.g., TKE or surface fluxes) is nudged — to demonstrate which change reduces the LWC/ $N_c$  bias most. The manuscript mentions sensitivity tests (not shown) — if you ran them, include a concise table/figure; if not, state explicitly that these are planned.

2. **Budget analysis is promising but hard to follow — improve clarity and quantitative presentation.**

Figures 11–12 and the text describe  $rc/N_c$  budgets, but the reader struggles to see magnitudes, relative contributions, and integrated tendencies. The text presents qualitative descriptions but few numbers. Add a compact table that reports integrated (over height or over the fog layer) mean tendencies for each budget term during each phase (formation, transition, adiabatic, dissipation). This will let the reader see which process dominates in each phase.

3. **Representation of the DSD: single-mode gamma assumption limits conclusions — quantify the impact and discuss alternatives.**

The text repeatedly notes that LIMA's single-mode/gamma DSD prevents realistic bimodality and under/overestimates large droplet population (e.g. missing 10  $\mu\text{m}$  mode or  $>30 \mu\text{m}$  tail). This is a central limitation that affects reflectivity and sedimentation conclusions. Add an explicit sensitivity test or offline exercise: take observed LWC and  $N_c$  and compute the gamma-law DSD (you have the "pink" distributions already). Then compute radar reflectivity and compare to observed reflectivity to quantify how much of the reflectivity bias is due to shape vs. other sources. Discuss possible microphysics remedies (multi-mode DSD, explicit coalescence parameterization, drizzle scheme) and estimate (qualitatively or quantitatively) which would be most effective.

4. **Instrumentation and representativeness caveats need clearer placement and stronger emphasis.**

The CDP measures 2–50  $\mu\text{m}$ ; the manuscript notes that droplets  $>50 \mu\text{m}$  could be important and that BASTA/CDP reflectivity mismatches suggest missing large drops.

But this important measurement limitation is discussed intermittently rather than in one explicit limitations paragraph; it is central to confidence in comparisons. Create a short dedicated subsection “Observational limits and impact on comparisons” that lists instrument ranges/uncertainties and then explicitly states how these affect each diagnostic (LWC,  $N_c$ , reflectivity). Provide a short table summarizing instrument vertical coverage, diameter ranges and uncertainties (you already have Table 1 but expand to explicitly link limitations to specific model comparisons).

5. **Phase definition thresholds and sensitivity — be explicit about robustness.**

Thin-to-thick transition is defined by several thresholds (TKE,  $dT/dz$ ,  $LW_{net}/LWP$ , CTH) and the manuscript notes sensitivity to averaging windows and layer definitions. This affects phase timing and subsequent aggregations. Provide a concise sensitivity test (e.g., show how transition start/end shift when using alternate thresholds or averaging windows) or at least a quantification of uncertainty ( $\pm$  hours). A short table or supplementary figure would suffice. This will strengthen confidence in phase-aggregated comparisons.

6. **Conclusions: strengthen practical takeaways for model developers/operational users.**

The Discussion lists useful suggestions (e.g., drizzle probes, radar Doppler) but the Conclusions are relatively general. The paper will be more impactful if it gives concrete, prioritized recommendations for model improvement and for observational follow-ups. Add a short bullet list in Conclusions: (1) short term model changes to try (e.g., adjusting deposition velocity, implementing two-mode DSD or drizzling scheme), (2) diagnostics to routinely output in future Meso-NH runs, and (3) highest-priority observations to collect in future campaigns.

## Minor

I give each minor point with a suggested edit or query. Page/line numbers are approximate (page numbers correspond to the PDF pages; lines are paragraph offsets).

1. **Abstract — be explicit about sample size and limitations.**

- Where: Abstract (p.1). add “two IOPs (IOP 6 and IOP 11) and N tethered-balloon profiles (state N if possible)” and a one-line note that aerosol data were missing for these dates (if true) so readers know sample size up front.

2. **Line in Intro: “To the authors’ knowledge, this is the first study...” — soften or justify.**

- Suggestion: Either cite LANFEX/LANFEX-like earlier tethered work explicitly (you do elsewhere) or replace with “one of the first” unless confident it is unique.

3. **Section 2.3 (Model set-up): clarify subgrid condensation treatment for 100 m vs 500 m runs.**

- p.7–8.
- Suggestion: briefly state whether any of the presented results are sensitive to the (non)use of subgrid condensation (e.g., was a test done to verify negligible impact at 100 m?). If not tested, state as caveat.

4. **Table 1 — make instrument uncertainties consistent and add CDP D-range explicitly in table column.**

- Table 1 (p.24).
- Suggestion: add a column “measurement diameter/range” and show that CDP is 2–50  $\mu\text{m}$  (already in text) so readers scanning the table see limitations.

5. **Figure captions: add sample counts (N profiles) and units.**
  - Figs. 7–10 (pp.32–36).
  - Suggestion: in each caption include “N = ... profiles aggregated” and ensure LWC units  $\text{g m}^{-3}$ ,  $\text{Nc cm}^{-3}$  and averaging window used.
6. **When referring to “excessive LWC” give typical bias numbers (e.g.,  $+0.15 \text{ g m}^{-3}$ ).**
  - results p.14–17.
  - Suggestion: quantify the bias ranges (mean  $\pm$  std) for each phase in a small table for clarity.
7. **Clarify the treatment of aerosol/radiative effects of aerosols.**
  - model setup/radiation (p.8).
  - Suggestion: you note that aerosol radiative effect is not included — briefly discuss whether including aerosol optical effect would be expected to change LWnet/LWP diagnostics.
8. **Grammar/typos: small issues to correct (examples):**
  - p.1 line “realisically” → “realistically” (check spelling throughout).
  - p.12 “thin,to thick” → “thin to thick” (remove stray comma).
9. **Make the “pink line” DSD exercise more visible (either move to main figures or add a small panel).**
  - DSD discussion (p.16–18).
  - Suggestion: place a small explanatory inset showing the observed vs gamma-reconstructed DSD and the consequences for reflectivity.
10. **Clarify how CCN/SMPS data were averaged and applied to activation parameterization.**
  - model set-up p.8.
  - Suggestion: state the time window used to produce the “mean activation spectra” and comment on representativeness when aerosol measurements were missing for the IOPs studied.
11. **In Section 3.2 (radiometer/lidar): discuss the 0.5 K cold bias of HATPRO (already mentioned elsewhere) and its impact on phase classification.**
  - p.9–11.
  - Suggestion: quantify whether shifting radiometer T by 0.5 K would change transition times.
12. **Line referencing: when you give times for IOP phases in Table 2, add UTC offset and explicit dates to avoid ambiguity for readers.**
  - Table 2 (p.24).
13. **Discussion: expand on whether LIMA’s activation parameterization (Vié et al., 2024) fully accounts for radiative cooling activation — you mention improvements but could be clearer.**
  - : Intro and model description (p.4–8).
14. **Data availability: the CDP files are “being finalized” — for reproducibility, give expected DOI or state that data will be provided upon acceptance.**
  - Data availability p.23–24.
15. **Reference to “Price, 2025” — verify final reference details (year 2025 cited for an in-prep source?)**
  - several places (Intro, etc.).
  - Suggestion: ensure all cited works are available or clearly marked as “in press”/“in review”.